

Into the unknown

Some scientific achievements are so impressive as to be nearly unbelievable. Quantum theory and the technological development of lasers, combined with Einstein's general theory of relativity, means that researchers using satellite-based atomic clocks accurate to one part in 10^{18} can now map gravitational anomalies closely enough to detect glaciers shifting under climate change. We're literally able to monitor the Earth by detecting changes in the local geometry of spacetime.

Yet, I'm often equally impressed by our astonishing ignorance in other areas. Until a few years ago, scientists were wholly ignorant even of the existence of *Prochlorococcus*, a simple cyanobacterium that turns out to be the most abundant photosynthetic organism on the planet, with some 10^{27} floating in the oceans. The global pandemic has made us all more aware of the family of coronaviruses, yet we know only about 5,000 distinct species of virus of any kind. This is despite expectations that the true number of species is in the many millions, if not billions. The vast world of viruses — reflecting some 10^{31} viral particles on Earth — remains largely unexplored.

In oceanic surveys a decade ago, for example, marine microbiologists sampled about 5,000 distinct viral populations from a variety of locations, but found that only 39 of those resembled lineages already known to science. More than 99% of the sampled viruses were hitherto unknown. In 2018, a study of a family of DNA viruses — known as the Megaviridae — found more genetic diversity within this one family than in all the bacteria and archaea combined (T. Mihara et al. *Microbes Environ.* **33**, 162–171; 2018).

Clearly, there's a viral world around us that we're barely aware of. Of course, with the events of the past year, we're suddenly much more aware of the threat of viral human pathogens and the coronaviruses in particular. Yet our knowledge of the full evolutionary tree of existing viruses of all types remains crude, even if it is now expanding rapidly due to the advent of so-called viral metagenomics — the study of genetic material sourced directly from the environment, foregoing the need for growing any cultures.

The narrow subset of RNA viruses illustrates the richness of this family tree.

This includes five major branches, each identified by a distinct RNA polymerase, the only protein shared by all RNA viruses. Each polymerase is the enzyme responsible for carrying out RNA transcription within that family. One distinct branch harbours the hepatitis C virus, another the poliovirus and the severe acute respiratory syndrome (SARS) coronavirus, as well as other coronaviruses. The influenza A virus inhabits another class (or phylum), as do the *Giardia lamblia* and Colorado tick fever viruses (J. Kuhn et al. *Nature* **566**, 318–320; 2019).



Learning more about what is probably the largest reservoir of genetic novelty on Earth can help us in trying to avoid global pandemics.

The evolutionary origin of the many branches of viral taxonomy dates back to the earliest era of life. Yet its connection to the tree of life remains somewhat paradoxical (H. M. B Harris & C. Hill *Front. Microbiol.* **11**, 3449; 2021). Viruses, in one sense, do not seem to be alive, as they cannot survive on their own. Their distinguishing characteristic is their lack of any ribosome, which means they cannot manufacture their own proteins. Rather, they hijack the cellular machinery of any host they infect to translate their messenger RNA into proteins. Hence, all viruses require the presence of other life to function.

Even so, this means that everything about viral biology is shaped by continuous evolutionary immersion in interaction with the living world — they're part of the living world even if only by their intense connection with it. As the US microbiologist Joshua Lederberg put it, "The very essence of the virus is its fundamental entanglement with the genetic and metabolic machinery of the host."

This explains why it is important for us to learn more about this viral world, which

harbours deep information about all life forms, including our own.

It's easy to forget how much of our basic knowledge of biology has come from the study of common microbes. Most antibiotics, for example, were first found in quite harmless soil bacteria. The staggering genetic diversity of the viral world almost certainly holds remarkable molecular components we currently know nothing about. The CRISPR revolution now transforming biology rests on discoveries linked to the way bacteria respond to viral infections, including viruses largely unknown to medical science. Soon, we may be able to harness viral capabilities for useful purposes — using viral infections against drug-resistant strains of bacteria, or discovering novel enzymes encoded by viruses useful in manufacturing new drugs.

And, of course, learning more about what is probably the largest reservoir of genetic novelty on Earth can help us in trying to avoid global pandemics. The threat that pathogens may jump from various hosts into humans is only going to grow as people continue to intrude on natural habitats and climate change induces species migrations. Filling in the taxonomy of existing viruses means identifying most of those that infect human tissues, as well as those that attack bacteria and other microbes within people, or in pets, pests and farm animals.

Yet efforts to build up detailed taxonomic knowledge of the viral world have, until fairly recently, been made by only a very small number of scientists. The effort started in 1886 with the classification of infectious agents of tobacco plants, and the later discovery that these agents passed through filters capable of removing all bacteria. A more systematic classification is now overseen by the International Committee on the Taxonomy of Viruses, which recognizes that most viruses will probably never be cultured, and that classification will come through faster genomic methods. Consequently, our current knowledge of viruses may be dwarfed by what we learn in the next few decades (see J. Kuhn *Encyclopedia Virol.* **1**, 28–37; 2021). □

Mark Buchanan

Published online: 6 October 2021
<https://doi.org/10.1038/s41567-021-01377-y>